OMICS A Journal of Integrative Biology
Volume 19, Number 7, 2015

DOI: 10.1089/omi.2015.0032

Towards New Drug Targets? Function Prediction of Putative Proteins of *Neisseria meningitidis* MC58 and Their Virulence Characterization

Mohd. Shahbaaz,[1] Krishna Bisetty,[1] Faizan Ahmad,[2] and Md. Imtaiyaz Hassan[2]

Abstract

Neisseria meningitidis is a Gram-negative aerobic diplococcus, responsible for a variety of meningococcal diseases. The genome of *N. meningitidis* MC58 is comprised of 2114 genes that are translated into 1953 proteins. The 698 genes (~35%) encode hypothetical proteins (HPs), because no experimental evidence of their biological functions are available. Analyses of these proteins are important to understand their functions in the metabolic networks and may lead to the discovery of novel drug targets against the infections caused by *N. meningitidis*. This study aimed at the identification and categorization of each HP present in the genome of *N. meningitidis* MC58 using computational tools. Functions of 363 proteins were predicted with high accuracy among the annotated set of HPs investigated. The reliably predicted 363 HPs were further grouped into 41 different classes of proteins, based on their possible roles in cellular processes such as metabolism, transport, and replication. Our studies revealed that 22 HPs may be involved in the pathogenesis caused by this microorganism. The top two HPs with highest virulence scores were subjected to molecular dynamics (MD) simulations to better understand their conformational behavior in a water environment. We also compared the MD simulation results with other virulent proteins present in *N. meningitidis*. This study broadens our understanding of the mechanistic pathways of pathogenesis, drug resistance, tolerance, and adaptability for host immune responses to *N. meningitidis*.

Introduction

Neisseria meningitidis MC58 belongs to the Gram-negative bacterial family of Neisseriaceae, and is an encapsulated and delicate aerobic diplococcus bacterium. Among the primary sources of bacterial meningitis worldwide, it is believed that only *N. meningitidis* can trigger epidemic conditions by causing manifestations such as pneumonia and sepsis (Jafri et al., 2013). This bacterium generally inhabits the cavity of the patient's nasopharynx. It causes life-threatening meningococcal diseases in children, especially in industrialized countries of Asia and Africa (Stephens et al., 2007). *N. meningitidis* is classified into 12 well-defined serogroups (viz., A, B, C, W, X, and Y) on the basis of the outcomes of genome typing techniques and their structural distinctions present in capsular polysaccharides, outer membrane proteins, as well as lipo-oligosaccharides (Broker et al., 2014; Rouphael and Stephens, 2012). Various recombinations and horizontal exchange of genes within the meningococcal genomes are responsible for the antigenic diversity among colonel complexes, leading to a distinctive clone expression (Read, 2014).

A variety of virulence factors are responsible for the pathogenesis of *N. meningitidis,* including capsular polysaccharides (CPS), lipo-oligosaccharide (LOS), and adhesins (Hill et al., 2010). The CPS, considered to be a key virulence factor, is comprised of N-acetyl-mannosamine-1-phosphate subunits with phosphodiester linkages (Fiebig et al., 2014). CPS enables this pathogen to escape the phagocytic complement-mediated mechanisms of the host and also forms the source for immunological serogrouping (Stephens et al., 2007). The sequenced genome of the MC58 strain of *N. meningitidis* contains 2,272,351 bps that are expressed in about 1953 proteins. It shows the presence of 2158 coding regions, of which 53.7% of the coding regions in the genome were allocated a biological function, while 35% were found to be ''hypothetical proteins (HPs)'' (Tettelin et al., 2000).

HPs are predicted proteins with no experimental validation at the biochemical level of protein expression (Shahbaaz et al., 2014a; 2014b). Approximately half of the proteins are not functionally characterized in the majority of the sequenced bacterial genomes. Identification of their natural functions will be useful in completing the available genomic information (Loewenstein et al., 2009; Nimrod et al., 2008).

[1]Department of Chemistry, Durban University of Technology, Durban, South Africa.
[2]Center for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, Jamia Nagar, New Delhi, 110025, India.

However, in order to understand the biochemical importance of these HPs, a precise functional annotation is necessary (Kumar et al., 2014a; 2014b). These predictions lead to the detection of new structure–function relationships and will be beneficial in uncovering an array of undetermined metabolic pathways in the organisms (Hassan and Ahmad, 2011; Hassan et al., 2008a; 2013b; Nimrod et al., 2008; Sinha et al., 2015; Zaidi et al., 2014). For structure-based drug design and discovery, we need some of the potential drug targets (Hassan et al., 2007a; 2007b; 2008b; 2009; 2010; 2013a; Naz et al., 2013; Thakur and Hassan, 2011; Thakur et al., 2013). Hence, HPs are also potential markers and targets for therapeutic agents in drug design and discovery (da Fonseca et al., 2012; Lubec et al., 2005; Minion et al., 2004).

We assumed that integration of several bioinformatics tools for sequence analyses and function predictions would be helpful in the annotations of these HPs (Shahbaaz et al., 2013). For this purpose, a sequence probing tool such as BLAST (Altschul et al., 1990) was used for searching the functional related homologs in various biological databases (Shahbaaz et al., 2013). Furthermore, tools such as Pfam (Bateman et al., 2002), SBASE (Vlahovicek et al., 2003), and SYSTERS (Krause et al., 2000) were used for the characterization of functional motifs and domains. These tools will provide a basis of identifying the role of these proteins in the biochemical processes (Rost and Valencia, 1996).

In a metabolically efficient biological network, the study of interactions between protein units and that of surrounding molecules by means of the STRING database (Szklarczyk et al., 2011) significantly improves the understanding of their roles in certain pathways. Finally, all above-mentioned tools were used to predict the functions of 698 HPs present in the genome of *N. meningitidis*. The functions of 363 HPs were predicted successfully. Among these proteins, the virulence factors were further classified using the information available from virulence prediction servers. The top two HPs with highest virulence scores were selected and subjected to Molecular Dynamics (MD) simulations to better understand their conformational behavior in a water environment.

Materials and Methods

The HPs present in the genome of *N. meningitidis* were annotated by using our previously designed protocols. This process of characterization contain three stages as depicted in Figure 1. The initial phase involved the characterization of HPs and their sequence retrieval from databases such as Uniprot (Apweiler et al., 2004). At the second level, HPs were subjected to various prediction tools concerning physicochemical properties, localization in subcellular environment, function, and domain annotation including interaction network. Finally, the precise functions were assigned to majority of HPs based on the consensus results obtained. Details of the methodology employed in each stage, followed by the MD simulations are presented below.

Sequence retrieval

The NCBI database (http://www.ncbi.nlm.nih.gov/genome/) provided a list of 698 HPs in the genome of *N. meningitidis*. The ''Gene ID'' of each HP was used to search the database of Uniprot (Apweiler et al., 2004), in order to obtain the sequence and primary accession number of these proteins. The sequences of HPs (<60 amino acids) were considered nonsignificant, since previous studies show that such proteins do not reveal the presence of any reliable homologs (Kumar et al., 2014b; Shahbaaz et al., 2013). Furthermore, if distinctive searches recognized a similar accession number for a group of HPs, then those proteins were considered as the ''redundant.'' The HPs fulfilling the above explained criteria were considered for further analyses.

Primary parameters' analyses

The Expasy's ProtParam server (http://web.expasy.org/protparam/) was used to predict the molecular weight, isoelectric point, and other physicochemical properties of non-redundant 670 HPs. Moreover, the prediction of the localization of these HPs in the cellular framework can be used to classify them in the form of targets to a vaccine that are generally membrane proteins (Hara et al., 2009). In the case of drug discovery, proteins localized in the cytoplasm are considered as the suitable targets (Yanamala et al., 2012). Since the localization studies of a majority of these HPs were not validated by the experimental studies, varieties of servers were used to localize the HPs appropriately. The PSLpred (Bhasin et al., 2005) uses support vector machine (SVM) based algorithms to predict the subcellular localization of bacterial proteins. These predictions were validated by the CELLO (Yu et al., 2004; 2006) and PSORTb (Yu et al., 2010) servers. Furthermore, SignalP 4.1 (Emanuelsson et al., 2007) was used to identify the presence of signal peptides in the sequences of HPs and SecretomeP (Bendtsen et al., 2005) to check the involvement of these proteins in nonclassical secretory pathways. Similarly, HPs involved in the mechanisms associated with transport were identified through the detection of transmembrane helices by TMHMM (Krogh et al., 2001) and HMMTOP (Tusnady and Simon, 2001) servers.

Function and domain analysis

The identification of sequence similarity in the biological databases (Altschul et al., 1990) was assumed to be a fundamental step for the function prediction of HPs (Shahbaaz et al., 2013). During the BLAST analysis, a protein with high sequence identity (>40%) with HPs was considered as an adjacent homolog, with the remaining as distant homologs. The sequence search that showed low identity between the sequences (≤20%) and a query coverage (<50%) were rejected. The results satisfying the above criterion of the sequence similarity searches were assumed to be the possible functional homologs of the respective HPs (see Supplementary Table S3).

Updated databases such as Pfam (Bateman et al., 2002), SYSTERS (Meinel et al., 2005), SMART (Letunic et al., 2012), Conserved Domain Database (CDD) (Marchler-Bauer et al., 2011), and SBASE (Vlahovicek et al., 2003) were used to identify the characteristic functional domains by utilizing the sequence of these HPs. CDD is an open source database present in the NCBI for the functional characterization of protein sequences by using annotation of conserved domain footprints. Similarly, SMART (Simple Modular Architecture Research Tool) was used to perform the search based on domain architecture and profiles in Swiss-Prot (Gasteiger et al., 2001), SP-TrEMBL (Bairoch and Apweiler, 2000), and

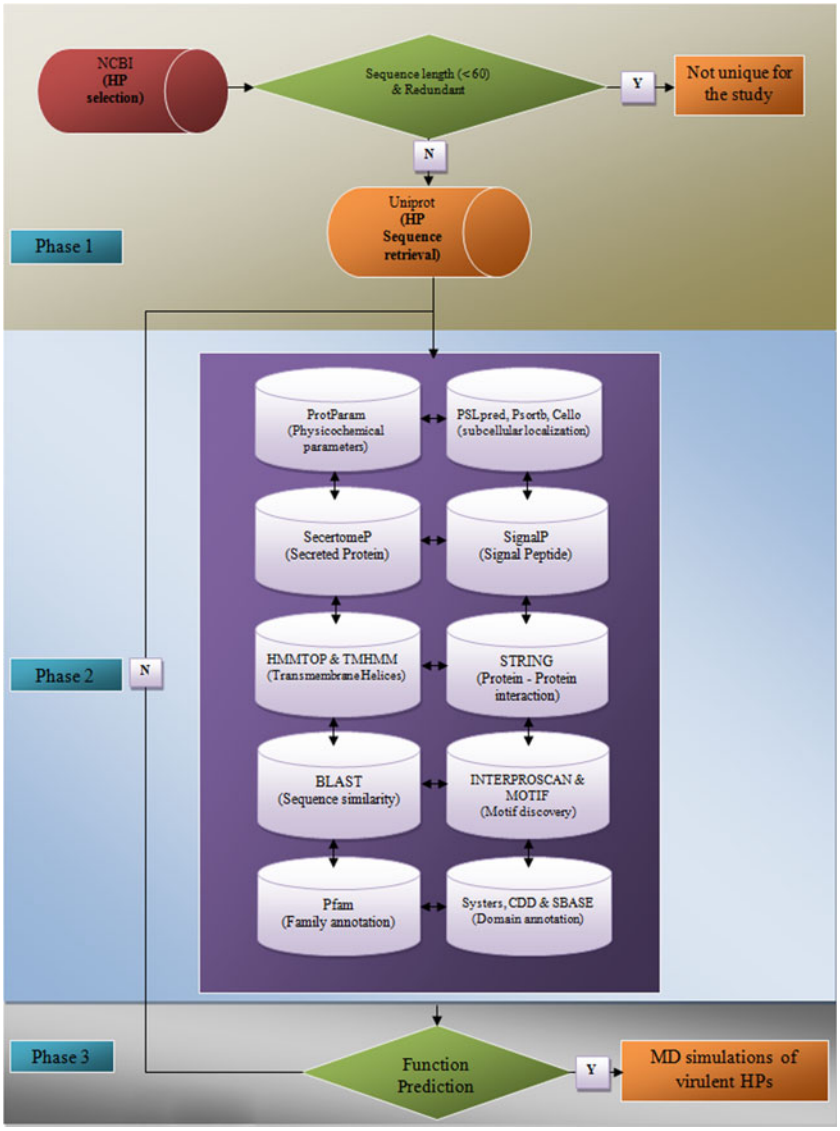

FIG. 1. The *in silico* approach used for annotating function of 698 HPs from *N. meningitidis.* The First Phase includes the characterization of HPs and selection of the protein for further studies. The Second Phase involves the calculation of various parameters using various available bioinformatics resources. Third Phase comprises the function prediction and selection of the virulent HPs. The selected virulence factors were subjected to the MD simulations to understand their structural behavior.

stable Ensembl (Hubbard et al., 2002) databases. In addition, the motif detection was performed using InterProScan (Quevillon et al., 2005) and MOTIF (http://www.genome.jp/tools/motif/) to predict the functional signature in the sequences. Further, in a metabolic network, adjacent proteins often transform the activity as well as the function of a given protein. Thus, information about the proceedings of the interactions between the proteins is assumed as crucial for predicting the precise function of a protein. Hence, we used STRING database (Szklarczyk et al., 2011) to predict functional partners of HPs in a biological network.

Virulence factor prediction

Virulence factors (VFs) are crucial for severity of disease (Baron and Coombes, 2007) and hence they are considered as effective targets in drug design and discovery. Varieties of un-annotated VFs were classified using VICMpred (Saha and Raghava, 2006) and VirulentPred (Garg and Gupta, 2008) servers that are SVM-based methods with an accuracy of 70.75% and 81.8%, respectively. Similarly, BTXpred (Saha and Raghava, 2007) and DBETH (Chakraborty et al., 2012) servers were used for prediction of the bacterial toxins.

Structure prediction

Three-dimensional (3-D) structure analyses were performed on the HPs with the highest predicted virulence scores. The BLAST module present in the Discovery Studio (DS) (Accelrys, 2013) was used for the identification of similar structural templates for the given HPs. The 3-D structures were predicted using the *ab initio* algorithm of I-TASSER

TABLE 1. PREDICTED ENZYMES AMONG THE GROUP OF HPS FROM GENOME OF *N. MENINGITIDIS**

S. No.	*Uniprot ID*	*Function*
		Transferase
1.	Q9K1J6	N-6 Adenine-specific DNA methylases
2.	Q9K196	Methyltransferase type 11
3.	Q9K0Z2	S-adenosyl-L-metheinine-dependent methyltransferase MidA
4.	Q9K0V4	S-adenosylmethionine-dependent methyltransferases (SAM or AdoMet-MTase),
5.	Q9JZ46	Methyltransferases
6.	Q9JZ42	Ribosomal RNA large subunit methyltransferase RlmN/Cfr
7.	Q7DDK6	Isoprenylcysteine carboxyl methyltransferase (ICMT)
8.	Q9JZH0	Ribosomal RNA large subunit methyltransferase J
9.	Q9JZ67	CheR methyltransferase, SAM binding domain
10.	Q9JYY8	S-adenosylmethionine-dependent methyltransferase
11.	Q9JYF4	SAM-dependent methyltransferase
12.	Q9JXS3	Uroporphyrin-III C/tetrapyrrole (Corrin/Porphyrin) methyltransferase
13.	Q9JXE3	RsmI AdoMet-dependent methyltransferase
14.	Q9JX96	Methyltransferase
15.	Q9K164	Glyscosyl transferase
16.	Q9K0Q3	Na+/H+ antiporter NhaC
17.	Q9K0H7	Bacterial transferase hexapeptide repeat
18.	Q9K047	Uracil phosphoribosyltransferase
19.	Q9JZX1	Acyltransferase
20.	Q9JZK0	Gcn5-related N-acetyltransferase (GNAT)
21.	Q9JZJ5	Glycine cleavage T protein (aminomethyltransferase)
22.	Q9JZG9	Glycerol-3-phosphate acyltransferase
23.	Q9JZ71	Nucleotidyltransferase
24.	Q9JY00	Putative glutamine amidotransferase type 2
25.	Q9JXU1	Rhodanese-related sulfurtransferases
26.	Q9JXP0	Aspartyl/glutamyl-tRNA amidotransferase
27.	P57090	Nicotinate-nucleotide adenylyltransferase
28.	Q9JXD5	Gcn5-related N-acetyltransferase (GNAT)
29.	Q9JXR9	Phosphoribosyltransferase
30.	Q9JXW5	Transposase
31.	Q9JYR5	Transposase IS200-like
32.	Q7DDH9	Bacteriophage transposase
33.	Q9K188	Aminoglycoside phosphotransferase
34.	Q9K0K2	CoA binding domain/acetyltransferase, GNAT family
35.	Q9K0A0	Na+/H+ antiporter
36.	Q9K071	rRNA methyltransferase
37.	Q9JRZ9	Carbohydrate kinase
38.	P57099	Glycerate kinase
39.	Q9JY37	Kinase (Couple_hipA, HipA_N, HipA_C)
		Peptidase
40.	Q9K1P6	Imelysin peptidase, Peptidase_M75
41.	Q9K1A1	Peptidase propeptide and YPEB domain protein
42.	Q9K163	Peptidase M23 family protein
43.	Q9K0U4	Peptidase propeptide and YPEB domain protein
44.	Q9JZF8	Peptidase M14, carboxypeptidase A
45.	Q9JZ20	Peptidase M23 family protein
46.	Q9JYU7	Peptidase S24 LexA
47.	Q4W566	Endopeptidase, NLPC/P60 domain
48.	Q9JY58	M61 glycyl aminopeptidase family protein
49.	Q9JY56	Peptidase M50
50.	Q9JYE3	Muramoyltetrapeptide carboxypeptidase
51.	Q9K0U1	L,D-transpeptidase catalytic domain
52.	Q9JXV3	Peptidase_M22
53.	Q9JS54	Peptidase S16
		Peroxidase
54.	Q9K1P5	Dyp-type peroxidase family
55.	Q9K036	DyP-type peroxidase
		Oxidoreductase
56.	Q9JRV5	Flavin containing amine oxidoreductase
57.	Q9K0A8	Multi-copper polyphenol oxidoreductase laccase
58.	Q9K157	FAD dependent oxidoreductase

(*continued*)

Table 1. (Continued)

S. No.	Uniprot ID	Function
59.	Q9K0Q5	Fe-S oxidoreductase of MoaA family
60.	Q9K077	FAD dependent oxidoreductase
61.	Q7DD95	Nitroreductase (YhbX/YhjW/YijP/YjdB) family protein
62.	Q7DDL4	Glutaredoxin (GRX) family
63.	Q9K178	Nitroreductase family protein
64.	Q9K184	Lysine decarboxylase
65.	Q9JZ90	Arsenate reductase
66.	Q9JYS5	Dioxygenase related to 2-nitropropane dioxygenase
67.	Q9JYH0	Carboxymuconolactone decarboxylase
68.	Q9JY86	4-Hydroxybenzoate octaprenyltransferase
69.	Q9JY11	NADPH-dependent FMN reductase
70.	Q9JY19	Hydroxylase (Cupin_4)
71.	Q9JYI1	Antibiotic biosynthesis monooxygenase
		Hydrolase
72.	Q9JXB2	Acyl hydrolase
73.	Q9JXE8	Serine hydrolase
74.	Q9JXK0	Murein hydrolase exporter LrgA/CidA family
75.	Q9JY57	Hydrolase
76.	Q9JY92	TatD like proteins, Predicted metal-dependent hydrolases with the TIM-barrel fold
77.	Q9JYB3	HAD hydrolase
78.	Q9JYE4	Hydrolase/nucleotide binding
79.	Q9JZG7	Nudix hydrolase domain
80.	Q9JZW7	Alpha/beta hydrolase family
81.	Q9K0P7	Predicted metal-dependent hydrolase
82.	Q9K0V2	ATP/GTP hydrolase
83.	Q9K137	Predicted hydrolase of the HAD superfamily
84.	Q9K191	Alpha/beta hydrolase fold family protein
85.	Q9K1D2	Allophanate hydrolase subunit 1
86.	Q7DDI9	Glycosyl hydrolase
87.	Q9JXJ9	Murein hydrolase transporter LrgB
		Lipase
88.	Q9JZ63	GDSL-like lipase
		Phosphatase
89.	Q9JZF6	Phosphoserine phosphatase
90.	Q9JXZ5	Polymerase/histidinol phosphatase-like
		Sulfatase
91.	Q9JYE0	Sulfatase-modifying factor enzyme 1
92.	Q9JYM7	Sulfatase
93.	Q7DD94	Sulfatase
94.	Q9JXJ7	Sulfatase
		Thioesterase
95.	Q7DD62	Thioesterase-like superfamily, Acyl-ACP thioesterase
96.	Q9JXZ1	Esterase family protein
97.	Q9K144	Peptidyl-prolyl cis-trans isomerase
		Translocase
98.	Q9JY32	Coronin binding protein
99.	Q9JYT7	Drug resistance translocase family protein
100.	Q9K0J7	Sec-independent protein translocase protein (TatC)
101.	Q9K0J6	Sec-independent protein translocase protein (TatC)
102.	Q9K0J1	Preprotein translocase YajC
		Synthase
103.	Q9JXA1	Thymidylate synthase
104.	Q9JXA4	tRNA threonylcarbamoyladenosine biosynthesis protein RimN
105.	Q9JXJ0	RNA pseudouridylate synthase family protein
106.	Q9JXP8	ATP synthase protein I
107.	Q9JY81	3-Oxoacyl-ACP synthase
108.	Q9JYN7	RNA pseudouridylate synthase
109.	Q9K1C4	Spermine/spermidine synthase
110.	Q9JYZ4	Rsu family of pseudouridine synthase signature
111.	Q9K020	Pseudouridine synthase, RsuA/RluB/E/F
		Ligase
112.	Q9JXD4	5-Formyltetrahydrofolate cyclo-ligase
113.	Q9K0K0	O-Antigen ligase family protein

(*continued*)

Table 1. (Continued)

S. No.	*Uniprot ID*	*Function*
		Restriction enzyme
114.	Q9JXE2	Predicted endonuclease distantly related to archaeal Holliday junction resolvase
115.	Q9JXH2	DNA recombination protein rmuC homolog
116.	Q9JXX1	Restriction endonuclease type II-like
117.	Q9JYB1	Nuclease
118.	Q7DD99	GIY-YIG nuclease superfamily
119.	Q9JYV3	Predicted 3'-5' exonuclease related to the exonuclease domain of PolB, RNase_H superfamily
120.	Q9JZJ4	Endoribonuclease L-PSP - like domain
121.	Q9JZK2	Smr protein/MutS2 Endonuclease activity
122.	Q9K050	TatD deoxyribonuclease family
123.	Q9K132	GYI-YIG nuclease superfamily
124.	Q9K1N6	GIY-YIG nuclease superfamily
		Metalloprotease
125.	Q9K1G9	Neutral zinc metallopeptidases
126.	Q9K1D4	Zinc metallopeptidases
127.	Q9JXI3	Zinc metalloprotease
		Hydrogenase
128.	Q9K0W7	Dehydrogenase (NnrS)
129.	Q9K0N5	Dehydrogenase (NnrS)
130.	Q7DDK1	Flavinator of succinate dehydrogenase
131.	Q9JZ97	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 3
132.	Q7DDA9	Glucose-6-phosphate 1-dehydrogenase family protein
133.	Q9JXK9	FmdE, Molybdenum formylmethanofuran dehydrogenase operon
		Lyase
134.	Q9JXT9	Diguanylate cyclase (GAF domain)
135.	Q9K028	Streptomyces cyclase/dehydrase - like domain
136.	Q9K084	Chorismate pyruvate lyase
137.	Q9K0S1	Lytic transglycosylase
138.	Q9K167	Adenylate cyclase (CYTH domain)
139.	Q9K0B6	Uracil DNA glycosylase like
		Nucleotidase
140.	Q9JZ83	ATPase
141.	Q9JXJ3	Cytosolic nucleotidase I
142.	Q9K080	P-loop ATPase protein family
143.	Q9K079	Ribosomal biogenesis GTPase
144.	Q9K0P1	ATPase (SMC)
145.	Q7DDP9	GTP-binding domain protein
146.	Q9K1B7	***GTPase, AAA ATPase domain***
		Carbonic Anhydrase
147.	Q7DD53	Carbonic anhydrase
		Phosphorylase
148.	Q9K0I1	Bifunctional kinase-pyrophosphorylase
149.	Q9K104	Nicotinate-nucleotide pyrophosphorylase
		Alanine racemase
150.	Q9K1N3	Alanine racemase, N-terminal
		Epimerase
151.	Q9K018	NAD dependent epimerase/dehydratase family protein
		Multifunctional enzymes
152.	Q9JZC1	RNses/ketopantoate reductase PanE/ApbA C terminal/ Protein phosphatase 2C-like - like domain;
153.	Q9JZE4	RNase/kinase/lyase

*The class of enzyme is indicated with Tan background.

(Roy et al., 2010). The assessment of the predicted models was carried out using TM score (Zhang and Skolnick, 2005) and Root Mean Square Deviation (RMSD) values computed using ITASSER. The selected models were subjected to energy minimization and optimization using the refinement modules of the DS. The reliability of structure predictions were analyzed by using the DALI server (Holm and Rosenstrom, 2010) that compare the generated 3-D structures of HPs with other proteins present in the structural databases.

Molecular dynamics simulation

Molecular dynamics (MD) simulation was performed on the modeled structures of the selected virulent HPs using

Table 2. HPs Involved in Transport Mechanisms from Genome of *N. meningitidis*

S. No	*Uniprot ID*	*Predicted functions*
		Transporters and carrier proteins
1.	Q9K1Q8	Integral membrane protein TerC
2.	Q9K1N9	Mechanosensitive ion channel MscS
3.	Q9K1M7	Oligopeptide transporter, OPT superfamily
4.	Q9K1J8	Major facilitator superfamily MFS_1 - like domain
5.	Q9K1E3	Mechanosensitive ion channel MscS
6.	Q7DDS2	Major Facilitator Superfamily domain, general substrate transporter
7.	Q9K166	Transporter
8.	Q9K149	Citrate transporter
9.	Q7DDQ9	Sugar ABC transporter substrate-binding protein
10.	Q9K0V6	Sulfite exporter TauE/SafE family protein
11.	Q9K0P8	Transporter associated domain, CBS domain
12.	Q9K0M2	Citrate transporter like domain
13.	Q9K0A5	RDD family transport protein
14.	Q9K099	Sodium: proton antiporter
15.	Q9K057	Drug/metabolite transporter
16.	Q9JZK6	Binding-protein-dependent transport systems inner membrane component
17.	Q9JZJ2	EamA-like transporter family
18.	Q9JZJ0	YceI-like domain, Outer membrane protein beta-barrel domain
19.	Q9JZG5	Cobalt uptake substrate-specific transmembrane region
20.	Q7DDH4	Curli production assembly/transport component CsgG
21.	P57062	FtsX-like permease family, MacB-like periplasmic core domain
22.	Q9JZ68	ABC-type transporter, periplasmic component
23.	Q9JZ54	Organic anion transporter polypeptide OATP
24.	Q7DDB9	Mn^{2+} and Fe^{2+} transporters of the NRAMP family
25.	Q9JYR8	Drug/metabolite transporter
26.	Q9JYH4	ABC transporter type 1, transmembrane domain
27.	Q9JY55	Ctr copper transporter family
28.	Q9JXY1	EamA-like transporter family protein
29.	Q7DD60	ABC transporter substrate-binding protein
30.	Q7DD59	ABC transporter, ATP-binding protein/Permease domain
31.	P67616	Fe(II) trafficking protein YggX
32.	Q9JXE9	RDD transporter protein family
33.	Q7DD46	ATP-binding cassette transporter nucleotide-binding domain
34.	Q9JXC4	Binding-protein-dependent transport systems inner membrane component
35.	Q9JXB6	Serine/threonine transporter SstT
36.	Q9JXN2	4-Amino-4-deoxy-L-arabinose transferase and related glycosyltransferases of PMT family
37.	Q9JYI6	Predicted permease YjgP/YjgQ family
38.	Q9JYI5	Permease YjgP/YjgQ, predicted
39.	Q9K1P7	Iron permease FTR1 family
40.	Q9K162	Inner membrane protein yhhQ/ Permease
41.	Q9K0X7	Xanthine/uracil/vitamin C permease
42.	Q9K0X4	Sulfite exporter TauE/SafE, Na+/Pi-cotransporter
43.	Q9K0W3	EamA-like transporter family
44.	Q9JZI1	CAP domain
45.	Q9JYG8	EamA-like transporter family, Permeases of the drug/metabolite transporter (DMT) superfamily
46.	Q9K1L1	MFS permease
47.	Q9K1K8	Positive regulator of sigma(E), RseC/MucC, Predicted permease YjgP/YjgQ family
48.	Q9K0Z7	Membrane fusogenic activity
49.	Q9K061	Cytochrome C assembly protein, transmembrane domain
50.	Q9K030	VanZ like family, Predicted membrane protein
51.	Q9JZW9	Outer membrane protein beta-barrel domain
52.	Q9JYS4	FixH inner membrane protein
53.	Q9JYP7	Integral membrane protein TerC
54.	Q9JS37	Major facilitator superfamily MFS_1 - like domain
55.	Q9JYJ1	Predicted membrane protein
56.	Q9JZN1	Membrane associated, eicosanoid/glutathione metabolism (MAPEG) protein
57.	Q9JZA0	RhaT l-rhamnose-proton symport 2
58.	Q9JY79	3-Hydroxylacyl-(acyl carrier protein)
59.	Q9JY61	PepSY-associated TM helix - like domain
60.	Q9JY52	Cytochrome c oxidase, subunit I
61.	Q9JXR5	Predicted membrane protein

(*continued*)

Table 2. (Continued)

S. No	*Uniprot ID*	*Predicted functions*
62.	Q9JXN7	ATPase, F0 complex
63.	Q9JXL5	Predicted membrane protein, Tripartite ATP-independent periplasmic transporters
64.	Q9JXE7	Transmembrane proteins
65.	Q9K0J9	Phosphate-selective porin O and P
66.	Q9JY87	asmA family Outer membrane protein
67.	Q9JRT4	Major surface glycoprotein MSG - like domain
68.	Q9JS17	Na^+/solute symporter
69.	Q9JZY4	Predicted periplasmic/secreted protein
70.	Q9JYH9	Electron transport protein SCO1/SenC
71.	Q9JXG2	Na^+/H^+ antiporter family, Predicted permease
72.	Q9JXN1	ABC-type transport system involved in resistance to organic solvents/ Toluene tolerance protein
73.	Q7DDR8	Predicted membrane protein, DoxX
74.	Q9K1L2	Transporter (MFS permease)
75.	Q7DDP8	H.8 outer membrane protein
76.	Q9K0Q1	Predicted membrane protein
77.	Q9JZV2	Pilus assembly protein PilW
78.	Q9JZV1	PilX N-terminal, Tfp pilus assembly protein PilX
79.	Q7DDJ8	Membrane protein

GROMACS (version 4.5.6; Van Der Spoel et al., 2005). In order to improve the electrostatic interactions, the virulent HPs were investigated with Optimized Potentials for Liquid Simulations–All Atoms (OPLS-AA) force-field and solvated in the single point charge (SPC) water model simulated using the Particle Mesh Ewald (PME) summation under periodic boundary conditions (PBC). An initial structure of the protein was energetically minimized with a convergence criterion of 0.005 kcal mol^{-1} using the steepest descent algorithm. The minimized structures were equilibrated for 1 ns using the NVT and NPT ensemble conditions. The MD simulations were performed for 100 ns using the LINCS algorithm of GROMACS with a time step of 2 fs. The g_cluster module in GROMACS was used for the clustering of generated conformations, based on the Javis Patrick clustering algorithm used with cutoff values set at 0.1 nm to organize the nearest structures. The central conformation present in the most densely populated cluster was used as the reference for the calculation of RMSD and Root Mean Square Fluctuation (RMSF). Further analyses were carried out using the utilities present in the GROMACS package.

Results

Outcomes of primary sequence analyses and function predictions are listed in Supplementary Tables S1–S4 (supplementary material is available online at www.liebertpub.com/omi). The consensus between the resulted prediction was formed on the basis of the similarities among the end results produced by tools in the adopted pipeline (Fig. 1). If three or more prediction tools detected or classified a similar function for any given HP, then it was assumed to be a possible function of the respective HP. These analyses showed that functions of 363 HPs from *N. meningitidis* were predicted successfully (Tables 1, 2, and 3). We further classified these annotated HPs among 41 functional categories of the proteins, which included 39 transferases, 14 peptidases, 16 oxidoreductases, 16 hydrolases, 4 sulfatases, 5 translocases, 9 synthase, 11 restriction enzymes, 6 hydrogenases, 6 lyases, 7 nucleotidases, 79 transporters and carrier proteins, 11 bacteriophage-related proteins, 20 virulence-related proteins, 10 immunity proteins, and 15 binding proteins (Fig. 2). These HPs showed high similarity to various functionally characterized proteins in the databases and may play an essential role in pathogenesis of the organism.

Sequences of 363 functionally annotated HPs were subjected to various virulence prediction servers. The VICMpred, based on the dataset of Gram-negative bacterial proteins, was assumed to be the primary method for the classification of virulent HPs. The classifications performed by VirulentPred, DBETH server, and BTXpred were used for the validation of VICMpred predictions (listed in Supplementary Table S5). The VICMpred classified 22 HPs as virulence factor (Table 4). The similarity search was performed for these proteins in genomes of other strains of *N. meningitidis* in order to identify the possible homologs (Table 5). The eight strains of *N. meningitidis* showed the presence of similar proteins in their proteome (Table 5). HP Q7DDQ9 and HP Q9JYT4 were predicted to be the most virulent proteins (Table 4).

In the absence of a reliable template in the protein data bank (PDB), structure of both proteins were predicted using the I-TASSER *ab intio* algorithm. The predicted model of HP Q7DDQ9 displayed 90.0% (i.e., 135) in the allowed region of the Ramachandran plot, while 8.7% of the residues belong to the marginal region and the rest in the disallowed region. Similarly, the 3-D structure of HP Q9JYT4 showed 88.9% in the allowed region, 9.2% in the marginal, and 1.9% in the disallowed region. Furthermore, the reliability of the predicted structures of both proteins were assessed using structural comparison protocol of DALI server. HP Q7DDQ9 showed structural similarities to lipopolysaccharide export systems with an RMSD range of 0.6–2.3 Å.

Similarly, HP Q9JYT4 demonstrated similarities to G-protein-signaling modulator 2 with RMSD ranges of 1.1–3.5 Å. The structure of HP Q7DDQ9 (Fig. 3A) was predicted using periplasmic lipopolysaccharide transport protein LptA (PDB ID–2R19). This protein showed a structure similar to lipopolysaccharide transport protein, with the lipopolysaccharides being the major class of *N. meningitidis* virulence factors produced by a robust pro-inflammatory reaction in the mammalian host during the meningococcal infection

Table 3. HPs Involved in Various Cellular Processes from Genome of *N. meningitidis*

S. No	*Uniprot ID*	*Predicted function*
		Lipoprotein
1.	Q7DD35	Lipoprotein
2.	Q9JXV4	Lipoprotein GNA1870-related, C-terminaL
3.	Q9JZQ7	Outer membrane lipoprotein
4.	Q9JZT7	YnbE-like lipoprotein
5.	Q9K1A0	Rare lipoprotein A (RlpA)-like double-psi beta-barrel, Lipoproteins
6.	Q9K1P8	Lipin, N-terminal conserved region
7.	Q9JZR5	Lipoprotein-34
		Transcriptional proteins
8.	Q9K1J7	Helix-turn-helix XRE-family like proteins
9.	Q9K146	YCII-related domain
10.	Q9K0W9	Rrf2 family transcriptional regulator
11.	Q9JZM5	mor transcription activator family protein
12.	Q9JZL3	IclR family transcriptional regulator
13.	P0A0Z0	Transcription regulator Rrf2-like
14.	Q9JYT8	Cysteine-rich domain
15.	Q9JYC7	Transcriptional regulator TACO1-like
16.	Q7DD45	XRE family transcriptional regulator/ Helix-turn-helix XRE-family like proteins
17.	Q9K1M3	Transcription elongation factor, N-terminal/ MafB protein
18.	Q9JXG7	Leucine zipper, homeobox-associated transcriptional factor
		Ribosomal protein
19.	Q7DD54	Iojap-like ribosome-associated protein
20.	P67217	Ribosomal maturation factor (Rim)
21.	Q9JZZ2	Ribosome associated, YjgA
		Bacteriophage related proteins
22.	Q7DDP4	Phage associated TspB-related protein
23.	Q9JZL2	BcepMu gp16 family phage-associated protein
24.	Q9JZE3	Phage-related protein
25.	Q9JZE2	Mu-like prophage protein gp29
26.	Q9JZD8	Phage minor structural protein GP20,
27.	Q9JZD6	Mu bacteriophage protein gp37
28.	Q9JYB9	Phage Mu protein F like protein
29.	Q9JZE1	Phage head morphogenesis, SPP1 gp7 family domain protein
30.	Q9JZD7	Bacteriophage Mu Gp36
31.	Q9JZC7	Bacteriophage Mu Gp46
32.	Q9JZC6	CheW-like protein - like domain
		Replication protein
33.	Q9JYK2	Replication initiation factor
34.	Q9JYD6	Replication initiation factor
		Translation protein
35.	Q9K0X0	Modification in translational fidelity (SUA5/yciO/yrdC)
		Ubiquitin protein
36.	P67259	Ubiquitin protein
		Structural motifs
37.	Q9K194	Gliding motility protein
38.	Q7DDP7	Roadblock/LC7 domain protein
39.	Q9JZG0	yecA family protein, SEC-C motif
40.	P64161	Septum formation initiator - like domain
41.	Q9K1K7	Peptidoglycan binding Lysin domain (LysM)
42.	Q9JZY3	MORN motif family protein
		Iron binding protein
43.	Q9JYT6	4Fe-4S ferredoxin-type iron-sulfur binding region signature
		Bax −1 inhibitor protein
44.	P63702	inhibitor of apoptosis-promoting Bax1 family protein
		Virulence-related protein
45.	Q9K1D3	Natural resistance-associated macrophage like
46.	Q9K0W8	Hemerythrin HHE cation binding region
47.	Q9JXB7	Surface antigen, Surface antigen variable number repeat
48.	Q9JY28	Pre-toxin domain with VENN motif family protein/hemagglutinin/hemolysin-related protein
49.	Q9JY47	Zonular occludens toxin (Zot), ATPases associated with a variety of cellular activities
50.	Q9JRY6	toxigenic (Zot)
51.	Q9JZ01	Multiple antibiotic resistance (MarC)-related proteins
52.	Q9JZV7	Viral DNA injection

(*continued*)

TABLE 3. (CONTINUED)

S. No	*Uniprot ID*	*Predicted function*
53.	Q7DDL0	Virulence protein
54.	Q9K0F2	Hedgehog/Intein domain,, Pretoxin HINT domain
55.	Q9K0S0	Pre-toxin domain with VENN motif
56.	Q9K0R9	VENN motif containing domain
57.	Q9K0S5	Pre-toxin domain with VENN motif/Hemagglutinin/hemolysin-related protein
58.	Q9K0S2	Pre-toxin domain with VENN motif family protein
59.	Q9K0S6	Possible hemagglutinin, Pre-toxin domain with VENN motif
60.	Q9K122	Pretoxin HINT domain, Putative toxin 46
61.	Q9K125	Pretoxin HINT domain
62.	Q9K1G5	Fusaric acid resistance protein family
63.	Q7DD57	Viral protein (NLPC/P60)
64.	Q9JY27	Hemagglutinin/hemolysin-related protein
		Tetratricopeptide-related protein
65.	Q9K165	Tetratricopeptide TPR_2 - like domain
66.	Q9K0T7	Tetratricopeptide like helical
67.	Q9K043	Heme-biosynthesis associated TPR protein
68.	Q9JZY0	Tetratricopeptide like helical
69.	Q9JZW5	TPR repeat family protein
70.	Q9JYT4	Tetratricopeptide repeat family protein
71.	Q9JYC1	MORN repeat variant
72.	Q7DDH5	Tetratricopeptide TPR_2 - like domain
		Cupin protein
73.	Q9JXU4	Cupin domain, 3-hydroxyanthranilic acid dioxygenase
		Lipopolysaccharide associated protein
74.	Q9K136	Lipopolysaccharide-assembly, LptC-related family protein
		Immunity protein
75.	Q7DD43	Immunity proteins (SMI1/KNR4 family)
76.	Q7DD41	Immunity protein 21
77.	Q7DDE0	Immunity protein 23, BNR/Asp-box repeat
78.	Q9K0R4	Immunity protein 29
79.	Q9K124	Immunity protein 22
80.	Q9K123	Immunity protein 22/MafB-related protein
81.	Q9K127	Immunity protein 22, Zn-finger in Ran binding protein
82.	Q9K140	Immunity protein 47
83.	Q9K1C8	Immunity proteins (SMI1_KNR4)
84.	Q9K1C5	Immunity protein 47
		Regulatory protein
85.	Q9JXB1	Regulatory/in cell division (Fic/DOC family)
86.	Q7DD48	Cell division protein (ZapA)
87.	Q7DD58	Regulatory mechanism of cell division
88.	Q9JY39	Regulatory/in cell division (Fic/DOC family)
89.	Q9JYU8	Regulatory/in cell division
90.	Q9JZ25	Regulatory Protein (Sel1)
91.	Q9K0V1	Regulatory mechanism of cell division (Fic/DOC)
		Binding proteins
92.	Q9K0E0	Nucleic acid binding
93.	Q9K0D2	Centromeric DNA binding (cnp1)
94.	Q9K074	Flavin-nucleotide-binding protein
95.	Q9K026	RNA binding (CRS1/YhbY)
96.	Q9K008	DNA-binding (NUMOD1)
97.	Q9JZU4	mRNA interferase PemK-like DNA binding protein
98.	Q9JZU1	KilA-N DNA binding protein
99.	Q7DDG5	S4 RNA-binding domain profile
100.	Q9JYT1	YbaB/EbfC DNA-binding family, EAP30/Vps36 family
101.	Q9JYD1	GTP-binding protein ERG
102.	Q9JY98	RNA binding/Osmoprotectant transporter
103.	Q9JXZ3	RNA-binding S4 domain
104.	Q9JXV0	RNA binding zinc-ribbon domain
105.	Q9JYZ0	Hemerythrin HHE cation binding domain
106.	Q9JRY5	Bacteriophage T5, Orf172 DNA-binding
		Miscellaneous proteins
107.	Q7DDU2	Competence damage-inducible protein A (CinA)
108.	Q9K1N5	Fusaric acid resistance protein-like (inner membrane)
109.	Q9K1K4	Signal peptide protein

(*continued*)

TABLE 3. (CONTINUED)

S. No	*Uniprot ID*	*Predicted function*
110.	Q9K182	Resistance protein TerB
111.	Q9K0Z5	Sporulation related domain (SPOR)
112.	Q9K0Y5	Signal transduction (SEL1)
113.	Q7DDQ0	Signaling protein
114.	Q9K0E2	Cell division membrane protein
115.	Q9K0A4	Stationary phase survival protein
116.	Q9K072	Gene expression up-regulator (SirB)
117.	Q9K019	Cofactor binding protein
118.	Q9JZQ6	Methionine biosynthesis (MetW)
119.	Q9JZJ3	Aminoacyl-tRNA editing
120.	Q9JS13	Capsule synthesis protein
121.	Q9JZA2	Membrane regulatory protein
122.	Q9JYN4	Universal stress protein
123.	Q9JYL3	Phosphatidylethanolamine-binding protein
124.	Q9JYL1	Cation-binding
125.	Q9JYJ0	Peroxiredoxin (OsmC)
126.	Q9JYJ5	Metal-binding protein
127.	Q9JYC2	Amino acid-binding (ACT)
128.	Q9JYB0	Antitoxin of toxin-antitoxin stability system
129.	Q9JXY3	Phosphoprotein
130.	Q9JXF2	Sporulation related domain
131.	Q9JRZ6	Ankyrin repeats mediate protein-protein interactions in very diverse families of proteins

(Zarantonelli et al., 2006). Therefore, this HP is considered as important for the pathogenesis. On the other hand, the 3D structure of HP Q9JYT4 showed all α-helix topologies (Fig. 4A) with a characteristic tetratricopeptide repeat (TPR).

In order to understand the conformational behavior of these HPs, we performed the comparison of various generated parameters during simulation processes. The HP Q7DDQ9 was found to be unstable as the RMSD plot showed a steep increase up to 25 ns (Fig. 3B). The radius of gyration showed a similar non-uniform behavior and the residues displayed higher fluctuations (Fig. 3C and D). The average RMSD, R_g, and total energy of the system were 0.43 nm, 1.68 nm, and -6.95×10^5 kJ mol^{-1}. Moreover, the MD simulations showed that HP Q9JYT4 is highly unstable,

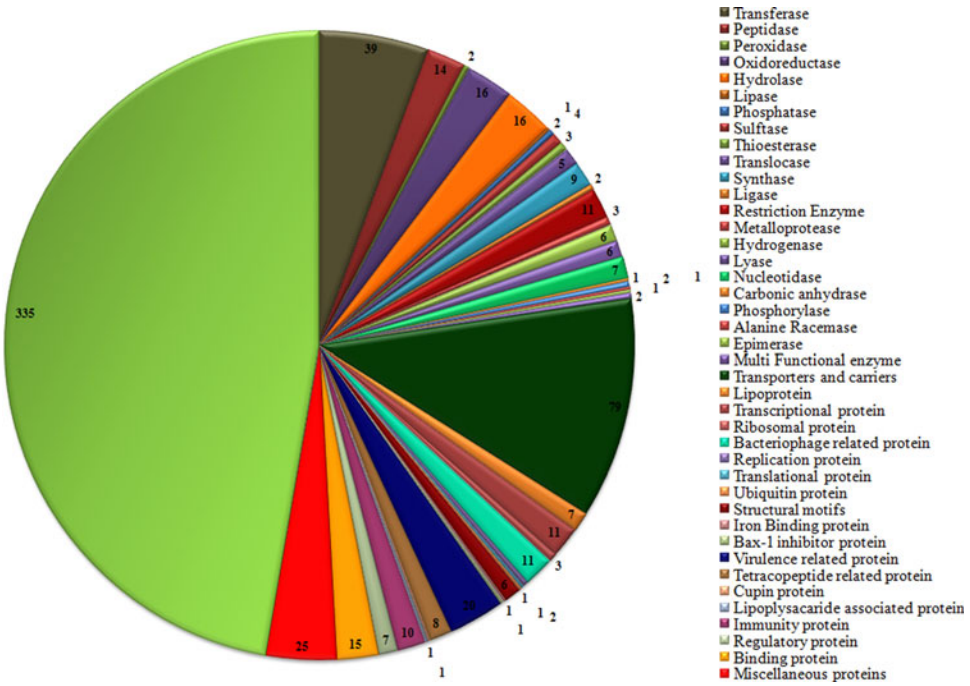

FIG. 2. The diagrammatic representation of functional categories identified in the set of 363 HPs from the genome of *N. meningitidis.*

Table 4. Predicted Virulence Factors and Toxins Proteins in the Set of 363 HPs

S.NO	*UNIPROT ID*	*VICMpred**	*VirulentPred**	*Toxin protein (DBETH server)*	*BTXpred Server: Prediction of bacterial toxins*
1.	Q9K1P6	1.31 (+)	(−) 0.89	(−)	Exotoxin
2.	Q9JZ20	1.35 (+)	(+) 1.02	(+)	(−)
3.	Q9JYE0	0.75 (+)	(−) 0.12	(+)	(−)
4.	Q9JY32	1.52 (+)	(+) 1.02	(+)	Exotoxin
5.	Q9JXH2	1.62 (+)	(−) 1.11	(−)	(−)
6.	Q9JXK9	1.47 (+)	(−) 1.05	(+)	(−)
7.	Q9K028	0.26 (+)	(−) 0.25	(+)	(−)
8.	Q7DDQ9	2.23 (+)	(+) 1.08	(+)	Exotoxin
9.	Q7DDH4	1.10 (+)	(+) 0.88	(−)	Exotoxin
10.	Q9K0X7	0.41 (+)	(−) 1.05	(−)	(−)
11.	Q9K0V2	0.67 (+)	(+) 0.92	(+)	Exotoxin
12.	Q9JYC7	1.21 (+)	(−) 0.99	(+)	Exotoxin
13.	Q9JZE2	0.91 (+)	(−) 1.08	(−)	(−)
14.	Q9JY28	0.17 (−)	(+) 0.63	(+)	Exotoxin
15.	Q9JY27	0.52 (−)	(+) 0.10	(+)	(−)
16.	Q9JRY6	0.66 (+)	(−) 0.58	(+)	(−)
17.	Q9K0R9	0.38 (+)	(+) 1.00	(−)	Exotoxin
18.	Q9K0P1	0.17 (−)	(+) 0.63	(+)	Exotoxin
19.	Q9K0S5	0.40 (+)	(+) 1.00	(+)	Exotoxin
20.	Q9K0S2	0.88 (−)	(−) 0.711	(+)	(−)
21.	Q9JYT4	1.63 (+)	(+) 1.09	(+)	Exotoxin
22.	Q9JXV0	0.99 (+)	(−) 1.04	(−)	(−)

Green, Most virulent; Yellow, protein classified as virulence factor with negative score.
*The Support Vector Machine (SVM) classification scores are presented in the form of results.

with continuous variations depicted by the RMSD and Rg plots throughout 100 ns (Fig. 4B and C). This result is complemented by the RMSF values indicating the presence of high energies in the constituent residues (Fig. 4D). On an average, the RMSD, R_g, and total energy values were 1.03 nm, 2.32 nm, and -1.23×10^6 kJ mol^{-1}, respectively. In comparison, their dynamic behavior was similar to the proteins obtained from VFDB [VFG0255 (Supplementary Fig. S1B–D) and VFG0260 (Supplementary Fig. S2B–D)]. These proteins showed continuous variations in the obtained parameters throughout 20 ns MD simulations. The RMSD values for these virulent proteins were found in the range of 0.4 nm–1.10 nm with total energy were between -1.44×10^6 kJ mol^{-1} to -4.40×10^6 kJ mol^{-1}. In addition, the R_g values fell among 2.35 nm–3.30 nm.

Discussion

With the exponential growth in genomic data, the number of protein sequences also increase rapidly in the biological databases (Gerlt et al., 2011). The proteins of unknown functions form the major part of the sequenced proteomes (Gerlt et al., 2011). The functional annotation of these proteins will allow precise understanding of the physiology and metabolism of various bacterial pathogens (Gerlt et al., 2011). Therefore, functional annotation of 698 HPs present in *N. meningitidis* will be significant in order to identify the chemotherapeutic targets against its pathogenesis. Due to limitations of the computational methods, only 363 HPs were annotated precisely. Furthermore, classification of protein into respective families is essential for phylogenetic analyses and functional characterization (Gerlt et al., 2011). Consequently we classified the 363 HPs into 41 classes of proteins. The identified functional classes are explained here in detail.

Enzymes

153 HPs were classified as enzymes (Table 1). Thirty-nine HPs with transferase-like properties were observed in the annotated set of proteins. The majority of this class showed SAM-dependent methyltransferase-like activities. These HPs may play an essential role in the process of methyl transfer to a variety of biomolecules such as proteins, nucleic acids, lipids, and small secondary metabolites (Struck et al., 2012). Hence, these HPs were considered significant for host–pathogen interactions (Struck et al., 2012). HP Q9JZ42 showed the presence of ribosomal RNA large subunit methyltransferase RlmN/Cfr activity (Toh et al., 2008) that may be responsible for alteration of adenosine in 23S rRNA, and may play an important role in the process of protein synthesis (Toh et al., 2008). Furthermore, HP Q9JXE3 was found to be a RsmI AdoMet-dependent methyltransferase. The latter enzyme is responsible for the maintenance of homeostasis by regulating the protein interactions and clearing the toxic substances (Lissina et al., 2013). The HP P57099 was identified as glycerate kinase, which may be involved in the phosphorylation of glycerate, particularly to 2-phosphoglycerate (Reher et al., 2006).

Fourteen HPs were categorized as peptidases. HP Q9K1P6 displayed imelysin-like peptidase activity and may be involved in iron uptake mechanisms (Xu et al., 2011). The muramoyltetrapeptide carboxypeptidase nature observed in HP Q9JYE3 may be essential for maintenance of the cell shape (DasGupta and Fan, 1979). Peroxidase-like activity was present in two HPs (Q9K1P5 and Q9K036). Therefore, we assumed that these antioxidant enzymes may be important for the survival of organism in the host (Moore and Sparling, 1995). Similarly, enzyme-like DsbA is among various significant oxidoreductase present in the periplasm of this

TABLE 5. PROTEINS BELONGING TO DIFFERENT STRAINS SHOWING SIMILARITY TO PREDICTED VIRULENT PROTEINS OF *N. MENINGITIDIS*

*Neisseria meningitidis strains (The predicted virulent proteins are shown with purple background)**															
S.NO	*MC58*	*Z2491*	*FAM18*	*053442*	*alpha14*	*8013*	*G2136*	*M01-240149*	*WUE 2594*	*alpha710*	*H44/76*	*M01-240355*	*M04-240196*	*NZ-05/33*	*510612*
1.	Q9K1P6	A1IPC4	A1KR62	A9LZT4	C6S9M8	C9X161	F0MCY2	F0MSV2	E7BIE3	E3D516	E6N0T2	F0MXP3	F0N2G4	-	-
2.	Q9JZ20	A1ISD7	A1KUF9	A9LZR1	C6S7H0	C9WZ07	-	-	E7BI28	E3D5K5	E6MZ93	-	-	-	-
3.	Q9JYE0	A1IT95	A1KV44	-	C6S882	C9WY96	-	-	E7BDU0	E3D1H4	E6N073	-	-	-	-
4.	Q9JY32	-	-	-	-	C9WXX1	-	-	-	E3D294	E6MW11	-	-	-	-
5.	Q9JXH2	Q9JWG3	A1KWC9	A9M094	C6S4M6	C9X2R7	-	-	E7BIV4	E3D3W0	E6MZN7	-	-	-	-
6.	Q9JXK9	A1IPR2	A1KW71	A9M0M7	C6S4T1	C9X2L4	-	-	E7BDI2	E3D3I7	E6MVH9	-	-	-	-
7.	Q9K028	A1IR45	A1KT55	A9M3M9	C6S662	C9X0B9	-	-	E7BG73	E3D3F5	E6N0D9	-	-	-	-
8.	Q7DDQ9	A1ITW0	A1KVS9	A9M2N5	C6S8V1	C9X2Y5	-	-	E7BF47	E3D173	E6MYT3	-	-	-	-
9.	Q7DDH4	A1IRW5	A1KTY3	A9LZ49	C6S6Z5	C9WZI1	-	-	E7BHK8	E3D4Q9	E6MWK7	-	-	-	-
10.	Q9K0X7	A1ITP2	A1KVK7	A9M2G7	C6S8N2	C9X3C8	-	-	E7BEQ2	E3D1K6	E6MX70	-	-	-	-
11.	Q9K0V2	A1ITL5	A1KVG9	A9M267	C6S8K6	C9X3F7	-	-	E7BEM7	E3D1N4	E6MUL4	-	-	-	-
12.	Q9JYC7	Q9JTB0	A1KV56	A9M1E0	C6S894	C9WY84	-	-	E7BE30	E3D1I8	E6MWW5	-	-	-	-
13.	Q9JZE2	A1IRV1	-	A9LZ31	-	C9WZJ9	-	-	E7BHI1	E3D4P0	F0MMF3	-	-	-	-
14.	Q9JY28	-	-	-	-	-	-	-	-	E3D1Z9	F0MJW5	-	-	-	-
15.	Q9JY27	A1IQB6	A1KSC4	A9M1X8	C6S5E0	C9X3J6	-	-	E7BER3	E3D1Z9	F0MJK3	-	-	-	-
16.	Q9JRY6	A1IT10	A1KR73	A9M0A6	-	-	-	-	-	E3D468	F0MKI8	-	-	-	-
17.	Q9K0R9	A1IQB4	A1KSC4	A9M1Y1	-	-	-	-	E7BER3	E3D2A0	E6MVK9	-	-	-	-
18.	Q9K0P1	A1IQE7	A1KSF4	A9M212	C6S5G6	C9X125	-	-	E7BEU9	E3D229	E6MVP5	-	-	-	-
19.	Q9K0S5	A1IQB6	A1KSC2	A9M1Y1	-	-	-	-	-	E3D2A0	F0MJW5	-	-	-	-
20.	Q9K0S2	A1IQB8	A1KSC6	A9M1Y1	C6S5E0	C9X3J6	-	-	E7BER3	E3D1S3	E6MYC7	-	-	-	-
21.	Q9JYT4	A1ISM8	A1KUP4	A9M0F8	C6S7R0	C9WYR4	-	-	E7BII4	E3D621	E6MXX5	-	-	-	-
22.	Q9JXV0	A1IQ25	A1KS24	A9M1F5	C6S538	C9X1Z9	-	-	E7BE25	E3D2R8	E6MV28	-	-	-	-

*Proteins with identity (<90%) are shown in *red color* while (>90%) in *black color*.

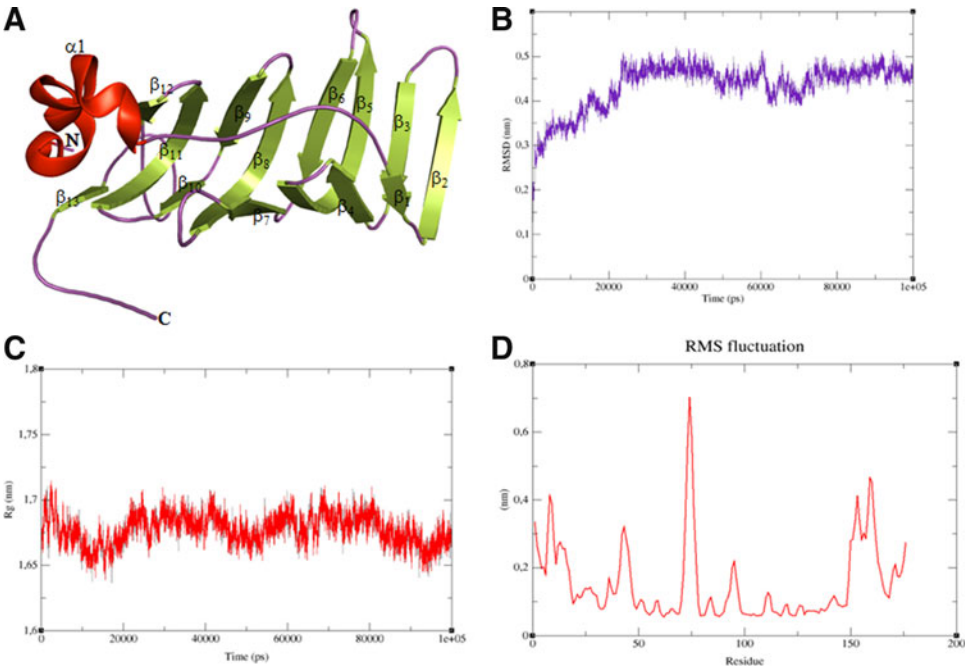

FIG. 3. **(A)** The predicted structure of HP Q7DDQ9. **(B)** The RMSD plot showing fluctuation during 100 ns MD simulation. **(C)** The plot of R_g against the time intervals indicating the unstable nature of the protein. **(D)** The RMSF of C^{α} atom for HP Q7DDQ9 showing high fluctuations of constituent residues.

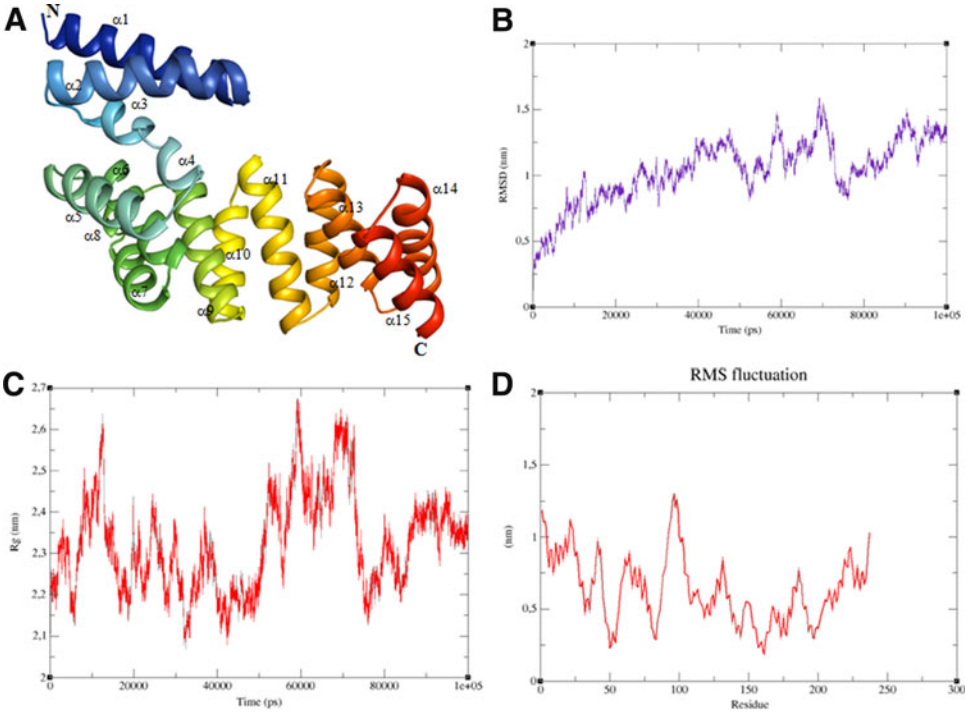

FIG. 4. **(A)** Predicted structure of HP Q9JYT4. **(B)** Plot of RMSD scores with sharp fluctuations during 100 ns MD simulation using explicit water conditions. **(C)** The plot of radius of gyration indicating that HP Q9JYT4 is unstable in nature. **(D)** RMSF plot of C^{α} atoms shows considerable higher fluctuations of the residues.

pathogen. These proteins are involved in the formation of the disulfide bond in various polypeptide chains for the dynamic folding of the membrane bound and surface exposed proteins (Vivian et al., 2009). We found 16 HPs possess oxidoreductase-like activity. HP Q9JY11 and HP Q9JZ90 showed the presence of NADPH-dependent FMN reductase and arsenate reductase activities, respectively. These enzymes may be a component of an arsenic resistance system, present in *N. meningitidis,* which may be responsible for the virulence (Neyt et al., 1997).

HP Q9JXE8 belongs to the family of serine hydrolases, which is a group of enzymes with many functions such as lipases and proteases (Kaschani et al., 2009). Moreover, HP Q9JXK0 and HP Q9JXJ9, which may be associated with the murein hydrolase activity, were grouped under 16 annotated hydolases and may be responsible for the separation of cells, architecture of membranes, and virulence of *N. meningitidis* (Adu-Bobie et al., 2004). HP Q7DDI9, predicted to be a glycosyl hydrolase, was considered to be involved in the hydrolysis of glycosidic bonds between the polysaccharides (Henrissat et al., 1995). Similarly the predicted sulfatase enzymes, associated with sulfur metabolism, were considered as potent drug targets because this process was found to be crucial for the pathogenesis of various bacterial species (Hatzios and Bertozzi, 2011). HP Q9K144 showed thioesterase activity with functionality related to peptidyl-prolyl cis-trans isomerase. The latter enzyme is involved in the process of protein folding and is significant for the virulence of pathogens by developing resistance against oxidative stresses (Reffuveille et al., 2012).

HP Q9JXA1 was predicted to be a thymidylate synthase. It may be involved in the catalytic conversion of deoxyuridine monophosphate into deoxythymidine monophosphate by the process of reductive methylation (Carreras and Santi, 1995). Furthermore, among the six characterized synthases, the HPs Q9JXJ0, Q9JYN7, Q9JYZ4 and Q9K020 were found to be similar to pseudouridine synthases. This group of enzymes is responsible for the post-transcriptional modifications of the RNA molecules by catalyzing site-specific isomerization of uridine residues (Hamma and Ferre-D'Amare, 2006).

The restriction enzyme of the *Neisseria* species triggers the damage of the DNA in the host cells during the infection, which affects the regulation of the cell cycle (Weyler et al., 2014). Due to this reason, the 11 HPs annotated as restriction enzymes were believed to be involved in the diseases caused by *N. meningitidis*. Among six characterized lyases, the HP Q9JXT9 was found to be similar to diguanylate cyclase, and it may be involved in the formation of cyclic nucleotides and hence may be responsible in regulating the virulence-associated mechanisms such as cytotoxicity (Kulasakara et al., 2006).

Chorismate pyruvate lyase-like activity was associated with HP Q9K084, indicating its involvement in the process of ubiquinone biosynthesis by the production of 4-hydroxybenzoate from chorismate (Nichols and Green, 1992). Moreover, lytic transglycosylases are the class of enzyme with similar substrate affinity as that of lysozyme that catalyzes the cleavage of peptidoglycans (Scheurwater et al., 2008). Therefore, HP Q9K0S1 may be involved in the maintenance of the bacterial cell structure. HP Q9K167 was an adenylate cyclase similar protein and may be involved in the conversion of adenosine triphosphate (ATP) into cyclic adenosine monophosphate (cAMP) (Steer, 1975). Among seven nucleases, we observed that HPs Q9JZ83, Q9K080, and Q9K0P1 may function as ATPases. These enzymes catalyze the breakdown of ATP into adenosine diphosphate (ADP) and free phosphate ions (Geider and Hoffmann-Berling, 1981). HP Q9K1N3 is homologous to alanine racemase and may be essential for the growth of *N. meningitidis* (Awasthy et al., 2012).

Transporter

The processes associated with transport such as nutrient uptake or waste product excretion are considered to be essential for the metabolism of the organism. Seventy-nine HPs were found to be associated with transport mechanisms (Table 2). These transporter and carrier proteins were generally associated with the survival and virulence of the disease-causing pathogens (Freeman et al., 2013). Annotation of seven HPs showed that they belong to the ATP-binding cassette (ABC 3) transporter family, possibly involved in the processes of the virulence because these HPs may be correlated with the uptake of iron, manganese, and zinc metal ions (Garmory and Titball, 2004). Furthermore, this class of protein is also associated with bacterial attachment to the mucosal surfaces of the host cells, which is crucial for pathogenesis. Therefore, these HPs are considered as the probable target (Garmory and Titball, 2004).

The phenomenon of multidrug resistance (MDR), distinguished by the instantaneous development of resistance against structurally and chemically diverse compounds, creates difficulties in treatment (Kumar and Varela, 2012). Active extrusion, which is one of the many causes of MDR, arises primarily due to the action of ABC and the major facilitator superfamily (MFS) transporters (Kumar and Varela, 2012). There are five HPs that may belong to the MFS superfamily of transporters and can be responsible for imparting the MDR in *N. meningitidis*.

HP Q9K149 is possibly involved in the transportation of citrate. Recent studies suggest that the intake of citrate may be essential for the virulence of Gram-negative bacteria (Urbany and Neuhaus, 2008). HP Q7DDB9 is member of the NRAMP family of Mn^{2+} and Fe^{2+} transporters. The uptake of manganese is also associated with bacterial pathogenesis (Papp-Wallace and Maguire, 2006). Similarly, HP Q9K1P7, characterized as the iron permease, may be involved in the transport of iron; uptake of this element plays a significant role in bacterial infection. A variety of bacterial virulence factors show the commencement of expression at lower levels of iron concentration (Braun, 2005). There is limited absorption of iron for bacterial pathogens within the host because of lower circulation of free iron (Braun, 2005). This form of iron leads to the formation of free radicals, which results in toxicity (Braun, 2005). Bacteria absorb iron by chelating proteins such as transferrin, heme, and lactoferrin (Braun, 2005). The member of family Neisseriaceae uses a siderophore-free mechanism for the uptake of iron from iron-binding proteins present in the host (Braun, 2005). These highly specific surface receptors are expressed by bacterial pathogens that cause direct removal of bound iron by coming in direct contact with the iron carrier proteins (Braun, 2005). Similarly, HP Q7DDB9 is member of the NRAMP family of Mn^{2+} and Fe^{2+} transporter; the uptake of manganese is also associated with bacterial pathogenesis (Papp-Wallace and Maguire, 2006).

Cellular processes

131 HPs were found to be involved in various cellular processes such as binding, transcription, translation, and replication (Table 3). There were 7 lipoproteins, 11 transcriptional-related HPs, 11 bacteriophage-related proteins, 6 structural motifs, 21 virulence-related HPs, and 16 binding proteins. The bacterial lipoproteins formed by modification of lipids causes the attachment of hydrophilic proteins to the hydrophobic surfaces by the means of hydrophobic interactions (Kovacs-Simon et al., 2011). This process of anchoring to the surface is important for many cellular processes, including the mechanism of virulence (Kovacs-Simon et al., 2011). There are seven HPs with functionality similar to lipoproteins, and therefore these proteins may play important roles in adhesion, alteration of inflammatory mechanisms, and virulence factors translocation (Kovacs-Simon et al. 2011). Eight HPs have a tetratricopeptide repeat (TPR); the activity of this structurally conserved motif is required for the assembly of multiple protein complexes. These motifs play significant roles in the virulence-associated cell processes (Kondo et al., 2010).

Proteins associated with the process of DNA-binding play an important role in bacterial pathogenesis. A structural motif such as the winged-helix-turn-helix in *Staphylococcus aureus* protein sarZ is associated with virulence by binding and activating *cvf* gene, which expresses into alpha hemolysin (Kaito et al., 2006). *Streptococcus* regulatory system, the virulence regulatory protein Srv, belongs to the family of CRP/FNR transcriptional regulators (Doern et al., 2008). The latter family members contain a characteristic helix-turn-helix motif (HTH) in the C-terminal region that is responsible for the DNA-binding (Doern et al., 2008). After mutation analysis of this motif, an alteration was observed in the protein–DNA interaction (Doern et al., 2008) that indicates its regulatory role in the virulence of bacteria. Furthermore, helix-turn-helix structural motifs are mainly present and associated with transcription regulation.

Transcriptional regulators such as HilC and HilD are involved in the DNA-binding processes of *Salmonella enterica*, leads to infection causing invasion of mammalian host cells (Olekhnovich and Kadner, 2002). Similarly, the HPs involved in RNA binding may also be involved in the pathogen survival in the host cells and also regulate its virulence factors (Ariyachet et al. 2013). Moreover, proteins like RfaH involved in the regulatory processes of *E. coli*, modulates the expression of other proteins and plays a major role in the virulence of Gram-negative bacteria (Nagy et al., 2002). Therefore these HPs of cellular processes may play an important role in the virulence of *N. meningitidis*.

MD simulation of virulent factors

Reliably predicted 3-D models of both HPs were subjected to 100 ns MD simulation. The solvated HPs were minimized at 1200 steps of steepest descent. Around 10,000 conformations were generated after the MD simulation. The generated conformations were clustered into 72 groups with an average RMSD of 0.1677 nm and 0.6710 nm for HP Q7DDQ9 and HP Q9JYT4, respectively. Furthermore, in order to understand the generated patterns for the respective HPs, we performed MD simulation on the virulence protein obtained from the VFDB database (Chen et al., 2005) for 20 ns. There were 43 proteins belonging to *N. meningitidis* among the dataset of 2454 bacterial virulence factor obtained from VFDB. The proteins VFG0255 and VFG0260 showed highest similarity to the respective virulent HPs were selected for MD simulations. VFG0255 is a capsule polysaccharide export outer membrane protein CtrA, while VFG0260 is a capsule polysaccharide modification protein LipB. The RMSD and R_g values for both the HPs were closely related to virulent proteins obtained from VFDB. The HP Q9JYT4 showed higher total energy while it was lower for HP Q7DDQ9. This closeness in the conformational behaviors indicated that these HPs may be involved in the virulence of *N. meningitidis*.

Conclusions

Of the 698 HPs investigated, 363 were successfully annotated by utilizing their sequences in the genome of *N. meningitidis*. There are many proteins whose function and virulence nature have been annotated in this study, which were initially undiscovered in the genome of this pathogen. These characterizations are complemented by the involvement of these proteins in the signaling as well as secretory pathways. Although the exact molecular function of an HP cannot be deduced from the *in silico* approaches, this framework will be helpful in the deduction of their molecular function based on the clustered conserved residues or the general fold characters.

In this study, the structure and dynamics of the characterized virulence factors were also assessed using molecular dynamics simulations in order to better understand their conformational behavior. Our study may be important in understanding the role of host–pathogen interaction and can further be utilized in the development of better therapeutic agents using HPs as a potential drug target.

Acknowledgments

The authors sincerely thank the Indian Council of Medical Research, Government of India for financial assistance (Project No. BIC/12(04)/2012). We express our gratitude towards the Centre for high performance computing, South Africa, and Jamia Millia Islamia, New Delhi for providing the computational infrastructure.

Author Disclosure Statement

The authors declare no conflict of interest regarding any financial and personal relationships with other people or organizations that could inappropriately influence (bias) this work.

References

Accelrys. (2013). Discovery Studio Modeling Environment, Release 3..5 (San Diego, Accelrys Software Inc.).

Adu-Bobie J, Lupetti P, Brunelli B, et al. (2004). GNA33 of Neisseria meningitidis is a lipoprotein required for cell separation, membrane architecture, and virulence. Infect Immun 72, 1914–1919.

Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. (1990). Basic local alignment search tool. J Mol Biol 215, 403–410.

Apweiler R, Bairoch A, Wu CH, et al. (2004). UniProt: The Universal Protein knowledge base. Nucleic Acids Res 32, D115–119.

Ariyachet C, Solis NV, Liu Y, Prasadarao NV, Filler SG, and McBride AE. (2013) SR-like RNA-binding protein Slr1 affects Candida albicans filamentation and virulence. Infect Immun 81, 1267–1276.

Awasthy D, Bharath S, Subbulakshmi V, and Sharma U. (2012). Alanine racemase mutants of Mycobacterium tuberculosis require D-alanine for growth and are defective for survival in macrophages and mice. Microbiology 158, 319–327.

Bairoch A, and Apweiler R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 28, 45–48.

Baron C, and Coombes B. (2007). Targeting bacterial secretion systems: Benefits of disarmament in the microcosm. Infect Disord Drug Targets 7, 19–27.

Bateman A, Birney E, Cerruti L, et al. (2002). The Pfam protein families database. Nucleic Acids Res 30, 276–280.

Bendtsen JD, Kiemer L, Fausboll A, and Brunak S. (2005). Non-classical protein secretion in bacteria. BMC Microbiol 5, 58.

Bhasin M, Garg A, and Raghava GP. (2005). PSLpred: Prediction of subcellular localization of bacterial proteins. Bioinformatics 21, 2522–2524.

Braun V. (2005). Bacterial iron transport related to virulence. Contrib Microbiol 12, 210–233.

Broker M, Bukovski S, Culic D, et al. (2014). Meningococcal serogroup Y emergence in Europe: High importance in some European regions in 2012. Hum Vaccin Immunother 10, 1725–1728.

Carreras CW, and Santi DV. (1995). The catalytic mechanism and structure of thymidylate synthase. Annu Rev Biochem 64, 721–762.

Chakraborty A, Ghosh S, Chowdhary G, Maulik U, and Chakrabarti S. (2012). DBETH: A database of bacterial exotoxins for human. Nucleic Acids Res 40, D615–620.

Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, and Jin Q. (2005). VFDB: A reference database for bacterial virulence factors. Nucleic Acids Res 33, D325–328.

da Fonseca MM, Zaha A, Caffarena ER, and Vasconcelos AT. (2012). Structure-based functional inference of hypothetical proteins from Mycoplasma hyopneumoniae. J Mol Model 18, 1917–1925.

DasGupta H, and Fan DP. (1979). Purification and characterization of a carboxypeptidase-transpeptidase of Bacillus megaterium acting on the tetrapeptide moiety of the peptidoglycan. J Biol Chem 254, 5672–5683.

Doern CD, Holder RC, and Reid SD. (2008). Point mutations within the streptococcal regulator of virulence (Srv) alter protein-DNA interactions and Srv function. Microbiology 154, 1998–2007.

Emanuelsson O, Brunak S, von Heijne G, and Nielsen H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2, 953–971.

Fiebig T, Freiberger F, Pinto V, et al. (2014). Molecular cloning and functional characterisation of components of the capsule biosynthesis complex of Neisseria meningitidis serogroup A: Towards in vitro vaccine production. J Biol Chem 289, 19395–19407.

Freeman ZN, Dorus S, and Waterfield NR. (2013). The KdpD/KdpE two-component system: integrating K(+) homeostasis and virulence. PLoS Pathog 9, e1003201.

Garg A, and Gupta D. (2008). VirulentPred: A SVM based prediction method for virulent proteins in bacterial pathogens. BMC Bioinformatics 9, 62.

Garmory HS, and Titball RW. (2004). ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. Infect Immun 72, 6757–6763.

Gasteiger E, Jung E, and Bairoch A. (2001). SWISS-PROT: Connecting biomolecular knowledge via a protein database. Curr Issues Mol Biol 3, 47–55.

Geider K, and Hoffmann-Berling H. (1981). Proteins controlling the helical structure of DNA. Annu Rev Biochem 50, 233–260.

Gerlt JA, Allen KN, Almo SC, et al. (2011). The Enzyme Function Initiative. Biochemistry 50, 9950–9962.

Hamma T, and Ferre-D'Amare AR. (2006). Pseudouridine synthases. Chem Biol 13, 1125–1135.

Hara Y, Mohamed R, and Nathan S. (2009). Immunogenic Burkholderia pseudomallei outer membrane proteins as potential candidate vaccine targets. PLoS One 4, e6496.

Hassan M, Kumar V, Somvanshi RK, Dey S, Singh TP, and Yadav S. (2007a). Structure-guided design of peptidic ligand for human prostate specific antigen. J Peptide Sci 13, 849–855.

Hassan MI, and Ahmad F. (2011). Structural diversity of class I MHC-like molecules and its implications in binding specificities. Adv Protein Chem Struct Biol 83, 223–270.

Hassan MI, Bilgrami S, Kumar V, Singh N, Yadav S, Kaur P, and Singh T. (2008a). Crystal structure of the novel complex formed between zinc α2-glycoprotein (ZAG) and prolactin-inducible protein (PIP) from human seminal plasma. J Mol Biol 384, 663–672.

Hassan MI, Kumar V, Singh TP, and Yadav S. (2007b). Structural model of human PSA: A target for prostate cancer therapy. Chem Biol Drug Design 70, 261–267.

Hassan MI, Shajee B, Waheed A, Ahmad F, and Sly WS. (2013a). Structure, function and applications of carbonic anhydrase isozymes. Bioorg Med Chem 21, 1570–1582.

Hassan MI, Toor A, and Ahmad F. (2010). Progastriscin: Structure, function, and its role in tumor progression. J Mol Cell Biol 2, 118–127.

Hassan MI, Waheed A, Grubb JH, Klei HE, Korolev S, and Sly WS. (2013b). High resolution crystal structure of human β-glucuronidase reveals structural basis of lysosome targeting. PloS One 8, e79687.

Hassan MI, Waheed A, Yadav S, Singh T, and Ahmad F. (2009). Prolactin inducible protein in cancer, fertility and immunoregulation: Structure, function and its clinical implications. Cell Mol Life Sci 66, 447–459.

Hassan MI, Waheed A, Yadav S, Singh TP, and Ahmad F. (2008b). Zinc α2-glycoprotein: A multidisciplinary protein. Mol Cancer Res 6, 892–906.

Hatzios SK, and Bertozzi CR. (2011). The regulation of sulfur metabolism in Mycobacterium tuberculosis. PLoS Pathog 7, e1002036.

Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, and Davies G. (1995). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. Proc Natl Acad Sci USA 92, 7090–7094.

Hill DJ, Griffiths NJ, Borodina E, and Virji M. (2010). Cellular and molecular biology of Neisseria meningitidis colonization and invasive disease. Clin Sci (Lond) 118, 547–564.

Holm L, and Rosenstrom P. (2010). Dali server: Conservation mapping in 3D. Nucleic Acids Res 38, W545–549.

Hubbard T, Barker D, Birney E, et al. (2002). The Ensembl genome database project. Nucleic Acids Res 30, 38–41.

Jafri RZ, Ali A, Messonnier NE, et al. (2013). Global epidemiology of invasive meningococcal disease. Popul Health Metr 11, 17.

Kaito C, Morishita D, Matsumoto Y, Kurokawa K, and Sekimizu K. (2006). Novel DNA binding protein SarZ contributes

to virulence in Staphylococcus aureus. Mol Microbiol 62, 1601–1617.

Kaschani F, Gu C, Niessen S, Hoover H, Cravatt BF, and van der Hoorn RA. (2009). Diversity of serine hydrolase activities of unchallenged and botrytis-infected Arabidopsis thaliana. Mol Cell Proteomics 8, 1082–1093.

Kondo Y, Ohara N, Sato K, et al. (2010). Tetratricopeptide repeat protein-associated proteins contribute to the virulence of Porphyromonas gingivalis. Infect Immun 78, 2846–2856.

Kovacs-Simon A, Titball RW, and Michell SL. (2011). Lipoproteins of bacterial pathogens. Infect Immun 79, 548–561.

Krause A, Stoye J, and Vingron M. (2000). The SYSTERS protein sequence cluster set. Nucleic Acids Res 28, 270–272.

Krogh A, Larsson B, von Heijne G, and Sonnhammer EL. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol 305, 567–580.

Kulasakara H, Lee V, Brencic A, et al. (2006). Analysis of Pseudomonas aeruginosa diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. Proc Natl Acad Sci USA 103, 2839–2844.

Kumar K, Prakash A, Anjum F, Islam A, Ahmad F, and Hassan MI. (2014a). Structure-based functional annotation of hypothetical proteins from Candida dubliniensis: A quest for potential drug targets. 3 Biotech 1–16.

Kumar K, Prakash A, Tasleem M, Islam A, Ahmad F, and Hassan MI. (2014b). Functional annotation of putative hypothetical proteins from Candida dubliniensis. Gene 543, 93–100.

Kumar S, and Varela MF. (2012). Biochemistry of bacterial multidrug efflux pumps. Int J Mol Sci 13, 4484–4495.

Letunic I, Doerks T, and Bork P. (2012). SMART 7: Recent updates to the protein domain annotation resource. Nucleic Acids Res 40, D302–305.

Lissina E, Weiss D, Young B, et al. (2013). A novel small molecule methyltransferase is important for virulence in Candida albicans. ACS Chem Biol 8, 2785–2793.

Loewenstein Y, Raimondo D, Redfern OC, et al. (2009). Protein function annotation by homology-based inference. Genome Biol 10, 207.

Lubec G, Afjehi-Sadat L, Yang JW, and John JP. (2005). Searching for hypothetical proteins: Theory and practice based upon original data and literature. Prog Neurobiol 77, 90–127.

Marchler-Bauer A, Lu S, Anderson JB, et al. (2011). CDD: A Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 39, D225–229.

Meinel T, Krause A, Luz H, Vingron M, and Staub E. (2005). The SYSTERS Protein Family Database in 2005. Nucleic Acids Res 33, D226–229.

Minion FC, Lefkowitz EJ, Madsen ML, Cleary BJ, Swartzell SM, and Mahairas GG. (2004). The genome sequence of Mycoplasma hyopneumoniae strain 232, the agent of swine mycoplasmosis. J Bacteriol 186, 7123–7133.

Moore TD, and Sparling PF. (1995). Isolation and identification of a glutathione peroxidase homolog gene, gpxA, present in Neisseria meningitidis but absent in Neisseria gonorrhoeae. Infect Immun 63, 1603–1607.

Nagy G, Dobrindt U, Schneider G, Khan AS, Hacker J, and Emody L. (2002). Loss of regulatory protein RfaH attenuates virulence of uropathogenic Escherichia coli. Infect Immun 70, 4406–4413.

Naz F, Anjum F, Islam A, Ahmad F, and Hassan MI. (2013). Microtubule affinity-regulating kinase 4: Structure, function, and regulation. Cell Biochem Biophys 67, 485–499.

Neyt C, Iriarte M, Thi VH, and Cornelis GR. (1997). Virulence and arsenic resistance in Yersiniae. J Bacteriol 179, 612–619.

Nichols BP, and Green JM. (1992). Cloning and sequencing of Escherichia coli ubiC and purification of chorismate lyase. J Bacteriol 174, 5309–5316.

Nimrod G, Schushan M, Steinberg DM, and Ben-Tal N. (2008). Detection of functionally important regions in ''hypothetical proteins'' of known structure. Structure 16, 1755–1763.

Olekhnovich IN, and Kadner RJ. (2002). DNA-binding activities of the HilC and HilD virulence regulatory proteins of Salmonella enterica serovar Typhimurium. J Bacteriol 184, 4148–4160.

Papp-Wallace KM, and Maguire ME. (2006). Manganese transport and the role of manganese in virulence. Annu Rev Microbiol 60, 187–209.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, and Lopez R. (2005). InterProScan: Protein domains identifier. Nucleic Acids Res 33, W116–120.

Read RC. (2014). Neisseria meningitidis; Clones, carriage, and disease. Clin Microbiol Infect 20, 391–395.

Reffuveille F, Connil N, Sanguinetti M, Posteraro B, Chevalier S, Auffray Y, and Rince A. (2012). Involvement of peptidylprolyl cis/trans isomerases in Enterococcus faecalis virulence. Infect Immun 80, 1728–1735.

Reher M, Bott M, and Schonheit P. (2006). Characterization of glycerate kinase (2-phosphoglycerate forming), a key enzyme of the nonphosphorylative Entner-Doudoroff pathway, from the thermoacidophilic euryarchaeon Picrophilus torridus. FEMS Microbiol Lett 259, 113–119.

Rost B, and Valencia A. (1996). Pitfalls of protein sequence analysis. Curr Opin Biotechnol 7, 457–461.

Rouphael NG, and Stephens DS. (2012). Neisseria meningitidis: Biology, microbiology, and epidemiology. Methods Mol Biol 799, 1–20.

Roy A, Kucukural A, and Zhang Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. Nat Protoc 5, 725–738.

Saha S, and Raghava GP. (2006). VICMpred: An SVM-based method for the prediction of functional proteins of Gram-negative bacteria using amino acid patterns and composition. Genomics Proteomics Bioinformatics 4, 42–47.

Saha S, and Raghava GP. (2007). BTXpred: Prediction of bacterial toxins. In Silico Biol 7, 405–412.

Scheurwater E, Reid CW, and Clarke AJ. (2008). Lytic transglycosylases: Bacterial space-making autolysins. Int J Biochem Cell Biol 40, 586–591.

Shahbaaz M, Ahmad F, and Hassan MI. (2014a). Structure-based function analysis of putative conserved proteins with isomerase activity from Haemophilus influenzae. 3 Biotech 1–23.

Shahbaaz M, Ahmad F, and Hassan MI. (2014b). Structure-based functional annotation of putative conserved proteins having lyase activity from Haemophilus influenzae. 3 Biotech 1–20.

Shahbaaz M, Hassan MI, and Ahmad F. (2013). Functional annotation of conserved hypothetical proteins from Haemophilus influenzae Rd KW20. PLoS One 8, e84263.

Sinha A, Ahmad F, and Hassan I. (2015). Structure based functional annotation of putative conserved proteins from Treponema pallidum: Search for a potential drug target. Lett Drug Design Discov 12, 46–59.

Steer ML. (1975). Adenyl cyclase. Ann Surg 182, 603–609.

Stephens DS, Greenwood B, and Brandtzaeg P. (2007). Epidemic meningitis, meningococcaemia, and Neisseria meningitidis. Lancet 369, 2196–2210.

Struck AW, Thompson ML, Wong LS, and Micklefield J. (2012). S-adenosyl-methionine-dependent methyltransferases: Highly versatile enzymes in biocatalysis, biosynthesis and other biotechnological applications. Chembiochem 13, 2642–2655.

Szklarczyk D, Franceschini A, Kuhn M, et al. (2011). The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 39, D561–568.

Tettelin H, Saunders NJ, Heidelberg J, et al. (2000). Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. Science 287, 1809–1815.

Thakur PK, and Hassan MI. (2011). Discovering a potent small molecule inhibitor for gankyrin using de novo drug design approach. Intl J Computat Biol Drug Design 4, 373–386.

Thakur PK, Kumar J, Ray D, Anjum F, and Hassan MI. (2013). Search of potential inhibitor against New Delhi metallo-beta-lactamase 1 from a series of antibacterial natural compounds. J Nat Sci Biol Med 4, 51.

Toh SM, Xiong L, Bae T, and Mankin AS. (2008). The methyltransferase YfgB/RlmN is responsible for modification of adenosine 2503 in 23S rRNA. RNA 14, 98–106.

Tusnady GE, and Simon I. (2001). The HMMTOP transmembrane topology prediction server. Bioinformatics 17, 849–850.

Urbany C, and Neuhaus HE. (2008). Citrate uptake into Pectobacterium atrosepticum is critical for bacterial virulence. Mol Plant Microbe Interact 21, 547–554.

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, and Berendsen HJ. (2005). GROMACS: Fast, flexible, and free. J Comput Chem 26, 1701–1718.

Vivian JP, Scoullar J, Rimmer K, et al. (2009). Structure and function of the oxidoreductase DsbA1 from Neisseria meningitidis. J Mol Biol 394, 931–943.

Vlahovicek K, Kajan L, Murvai J, Hegedus Z, and Pongor S. (2003). The SBASE domain sequence library, release 10: Domain architecture prediction. Nucleic Acids Res 31, 403–405.

Weyler L, Engelbrecht M, Mata Forsberg M, Brehwens K, Vare D, Vielfort K, Wojcik A, and Aro H. (2014). Restriction endonucleases from invasive Neisseria gonorrhoeae cause double-strand breaks and distort mitosis in epithelial cells during infection. PLoS One 9, e114208.

Xu Q, Rawlings ND, Farr CL, et al. (2011). Structural and sequence analysis of imelysin-like proteins implicated in bacterial iron uptake. PLoS One 6, e21875.

Yanamala N, Gardner E, Riciutti A, and Klein-Seetharaman J. (2012). The cytoplasmic rhodopsin-protein interface: Potential for drug discovery. Curr Drug Targets 13, 3–14.

Yu CS, Chen YC, Lu CH, and Hwang JK. (2006). Prediction of protein subcellular localization. Proteins 64, 643–651.

Yu CS, Lin CJ, and Hwang JK. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein Sci 13, 1402–1406.

Yu NY, Wagner JR, Laird MR, et al. (2010). PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26, 1608–1615.

Zaidi S, Hassan MI, Islam A, and Ahmad F. (2014). The role of key residues in structure, function, and stability of cytochrome-c. Cell Mol Life Sci 71, 229–255.

Zarantonelli ML, Huerre M, Taha MK, and Alonso JM. (2006). Differential role of lipooligosaccharide of Neisseria meningitidis in virulence and inflammatory response during respiratory infection in mice. Infect Immun 74, 5506–5512.

Zhang Y, and Skolnick J. (2005). TM-align: A protein structure alignment algorithm based on the TM-score. Nucleic Acids Res 33, 2302–2309.

Address correspondence to:
Md Imtaiyaz Hassan, PhD
Center for Interdisciplinary Research in Basic Sciences,
Jamia Millia Islamia
Jamia Nagar
New Delhi 110025
India

E-mail: mihassan@jmi.ac.in